



LUND UNIVERSITY
School of Economics and Management

Measuring Risk with Expected Shortfall

Comparison of Expected Shortfall and Value at Risk

by

Huan Liu

Stacy Kuntjoro

June 2015

Master's Programme in Finance

Supervisor: Birger Nilsson

Abstract

In 2012, The Basel Committee on Banking Supervision decided to change the standard risk measure from the well-known Value-at-Risk (*VaR*) to Expected Shortfall (*ES*). The committee believes that the new standard risk measure could offer more benefit, aside from just overcoming the major weaknesses of *VaR* like incoherency and inability to capture tail risk. In this study, best models for *VaR* and *ES* are determined and a comparison between the best models of the newly implemented risk measure and the former risk measure is amplified.

Four estimation approaches are used to estimate best models for *VaR* and *ES*: Historical Simulation, Gaussian distribution, Student t-distribution and Extreme Value Theory (Peak Over Threshold model). From these four approaches, nine models of estimation are developed: HS model, conditional and unconditional N-distribution, conditional and unconditional t-distribution, conditional EVT with $\xi = 0$ and $\xi \neq 0$, as well as unconditional EVT with $\xi = 0$ and $\xi \neq 0$.

The evaluation window for this research consists of three distinctive periods: bad economy period, recovery period and good economy period. These three periods are deliberately chosen to apprehend the impact of different economic situations on the models' ability to forecasts *VaR* and *ES*.

The final results of the research indicated that in general the models that work best for *VaR* are the same models that work best for *ES*. Amongst all nine models, the unconditional EVT model seems to be the only suitable model for all the three evaluation periods.

Keywords: *risk measure, expected shortfall, value at risk, Basel III agreement, tail risk, coherent, parametric approach, non-parametric approach.*

Acknowledgements

We would like to extend our deepest and greatest gratitude to these individuals who have helped and supported us throughout our research and made this manuscript possible:

Lund Global Scholarship, for enabling us to have this great opportunity to study in Lund.

Birger Nilsson, our thesis supervisor who is also a Financial Valuation and Risk Management specialist, for his valuable insights along with endless support, guidance and assistance in helping us finish this manuscript.

Hossein Asgharian, our programme director, for his endless guidance throughout the entire Master Programme course in Sweden.

Julienne Stewart-Sandgren, our educational developer and academic writing tutor, for her constant encouragement and assistance in overcoming difficulties related to the writing of the manuscript.

Rikke Barthélemy, our master coordinator, for her kind support and assistance in dealing with administrative issues.

Lecturers, Lab Assistance and Administrators, for making this programme possible and for sharing their time and immense knowledge with us.

Families and Friends, for believing and being there for us when we needed them most.

Huan & Stacy

Table of Contents

- 1 Introduction 7**
 - 1.1 Background 7
 - 1.2 Research Purpose 9
 - 1.3 Research Limitations 10
 - 1.4 Outline of the Thesis 11
- 2 Literature/Theoretical Review 12**
- 3 Methodology 18**
 - 3.1 Estimation Approach 18
 - 3.1.1 Historical Simulation (HS) 18
 - 3.1.2 Gaussian Distribution (N-dist) 19
 - 3.1.3 Student t-Distribution (t-dist) 20
 - 3.1.4 Extreme Value Theory (EVT) 21
 - 3.2 Model Evaluation 23
 - 3.2.1 Backtesting VaR: Kupiec Frequency Test 24
 - 3.2.2 Backtesting ES 25
 - 3.3 Data Analysis 27
- 4 Analysis and Discussion 29**
 - 4.1 ES Estimates and Backtesting 29
 - 4.2 VaR Estimates and Kupiec Tests 33
 - 4.3 Comparison of Best Models 37
- 5 Summary and Conclusion 39**
- References 41**
- Appendix A 47**
- Appendix B 48**
- Appendix C: Part A 49**
- Appendix C: Part B 50**
- Appendix C: Part C 51**
- Appendix D: Part A 52**
- Appendix D: Part B 53**
- Appendix D: Part C 54**

List of Tables

Table 1:5% and 0,01% significance thresholds for Z^2	27
Table 2:Loss observations: Descriptive Statistics	28
Table 3:Backtesting Results for Expected Shortfall	32
Table 4:Kupiec Backtest Results for Value-at-Risk	36
Table 5:Parameters estimates for Extreme Value Theory.....	48

List of Figures

Figure 1: The loss observations for the entire sample period 27

Figure 2: Expected Shortfall - HS 49

Figure 3: Expected Shortfall - N-dist 49

Figure 4: Expected Shortfall – t-dist 50

Figure 5: Expected Shortfall - Unconditional EVT 50

Figure 6: Expected Shortfall - Conditional EVT 51

Figure 7: Value at Risk - HS 52

Figure 8: Value at Risk - N-dist 52

Figure 9: Value at Risk - t-dist 53

Figure 10: Value at Risk - Unconditional EVT 53

Figure 11: Value at Risk - Conditional EVT 54

List of Abbreviations

AWHS	: Age Weighted Historical Simulation
BRW	: Boudoukh, Richardson and Whitelaw Historical Simulation
DQT	: Dynamic Quantile test
ES	: Expected Shortfall
EVT	: Extreme Value Theory
EVT-cond	: Conditional Extreme Value Theory
EWMA	: Exponentially Weighted Moving Average
FHS	: Filtered Historical Simulation
GARCH	: Generalized Auto-Regressive Conditional Heteroskedasticity
GPD	: Generalized Pareto distribution
HS	: Basic Historical Simulation
IG-CARES	: Indirect GARCH - Conditional Autogressive Regression
LM ARCH	: Long Memory Auto-Regressive Conditional Heteroskedasticity
MHS	: Mirrored Historical Simulation
N-dist	: Normal distribution
n.d.	: No date
POT	: Peak Over Threshold
RWA	: Risk Weighted Assets
t-dist	: Student t-distribution
VaR	: Value at Risk
VWHS	: Volatility Weighted Historical Simulation
VCV	: Normal Variance-Covariance Approach

1 Introduction

1.1 Background

Despite the traditional economic theory of the risk management irrelevant proposition; the financial market with all of its participants do need effective risk management system. Given the imperfections of the real financial market, its players have to be ready to face the worst-case scenario of losing their investment. This is especially important for financial institutions, which has the power to cause major economic breakdown.

The collapse of financial systems and/or institutions could provoke an economic crisis on a global scale. Such undesirable incidents have happened several times throughout the history. One of the major incidents occurred in 1973 after the disintegration of the Bretton Wood system, a multinational monetary management system in charge of supervising the exchange rates among world's major industrial states. This failure caused significant losses on foreign exchange and led many banks to default.

To anticipate the occurrence of similar financial market turmoil, the central bank governors of the G10 countries formed a supervisory committee in 1974, which then later named the Basel Committee on Banking Supervision. This committee aims to enhance financial stability by setting a series of international banking regulations to endorse common understanding and improve cross-border cooperation ([Basel Committee on Banking Supervision, 2014](#)).

The Basel Committee believes that having a strong international banking system could prevent another systemic breakdown and is therefore fundamental to the overall financial stability. Since the beginning of Basel I, the committee has established multiple sets of regulations that converged on the concept of capital adequacy, which is considered as one of the key determinants of strength in the banking system ([Basel Committee on Banking Supervision, 1988](#)). According to The Association for Financial Markets in Europe (AFME), capital “serves as a buffer to absorb unexpected losses” ([AFME, n.d.](#)). In confirmation with AFME, the Reserve Bank of New Zealand also stated that capital could be used to “absorb a

reasonable level of losses before becoming insolvent” ([Reserve Bank of New Zealand, 2007](#)). Retrospectively, establishing capital adequacy could also be viewed as a risk management tool that could enhance financial stability and efficiency.

The prerequisite amount of loss absorbing fund a bank or financial institution has to own in order to be considered by the financial regulators as having adequate capital is known as the capital requirement. According to the [1996 Basel Amendment](#), banks and other financial institutions will be charged with capital requirement incidental to the amount of market risk and credit risk they are exposed to. Under the Basel rules, the minimum capital requirement for banks and financial institutions is set to be 8% of their so called risk weighted asset (RWA), which is a measure of the bank’s total credit exposure ([Hull, 2012](#)).

In 1996, the Basel Accord Amendment introduced Value at Risk (abbreviated as *VaR*) as the standard risk measure as well as the quantitative risk metrics used to determine the regulatory capital requirements. In his book, *Risk Management and Financial Institution*, [Hull \(2012\)](#) defines *VaR* as “an attempt to provide a single number that summarizes the total risk in a portfolio” (p.183). In addition, [Acerbi and Szekely \(2015\)](#) equivalently defined *VaR* as “the best of worst $x\%$ losses” (p.4). As a risk measure, *VaR* could be used to determine the RWA, the necessary variable to calculate the minimum capital requirement. Unfortunately, due to the several major weaknesses found in *VaR*, incoherency and inability to capture tail risk, the committee has agreed to use an alternative measure to calibrate risk along with the capital requirement ([Kerkhof & Melenberg, 2004](#)).

In 2012, the Basel Committee on Banking Supervision issued a notion to replace *VaR* with Expected Shortfall (abbreviated as *ES*) through its first consultative documents for *Fundamental Review of The Trading Book*. [Hull \(2012\)](#) defined *ES* as “the expected loss during time T conditional on the loss being greater than the X^{th} percentile of the loss distribution” (p.187). In other words, *ES* are the average losses that are larger than *VaR*, which implies it has the ability to capture tail risk. In addition to its ability to capture tail risk, *ES* also offers other advantages such as subadditivity, mathematical tractability, uniqueness, and uses same the risk models ([Acerbi & Szekely, 2015](#)).

Unfortunately, this new viable risk measure also brings forth a theoretical debate: for many years, market participants have assumed that *ES* cannot be backtested. [Acerbi & Szekely \(2015\)](#) stated that backtesting is a test used to ensure that the model yields forecasts that are in

line with the actual realisations. In other words, it is used to check the validity and the reliability of the model in capturing risk. Therefore, in order to discover the best model for expected shortfall, an appropriate backtesting model should also be defined. Unlike its preceding risk metrics (*VaR*), the backtesting approach for *ES* is still underdeveloped due to the claim that it is not elicitable. To put it bluntly, the committee is now facing a dilemma of having a subadditive yet tail sensitive measure, but might not be robust ([Ziegel, 2014](#)). To overcome this problem, in 2013 the Basel Committee agreed to compromise by continuing to adopt the *ES* approach while keeping the backtesting approach for *VaR* ([MSCI Inc., 2014](#)).

The Basel Committee on Banking Supervision has planned to fully implement the Basel III Accord in 2019 ([Saigol & Fleming, 2013](#)). This means that by 2019 banks and other financial institutions will fully adopt the new risk measure, *ES*. Given the fairly short period to prepare for the transition, the need to develop a backtesting methodology for *ES* has grown significantly. Luckily, one of the leading providers of investment decision support tools, MSCI, has recently announced a breakthrough in the backtesting methodology for *ES*. [Acerbi and Szekely \(2014\)](#), the Executive Director and Senior Researcher of MSCI, have successfully developed three sophisticated methodologies to backtest *ES*. These methods are available for public through their publication called *Backtesting Expected Shortfall*.

1.2 Research Purpose

Since *ES* will soon replace *VaR* as the standard risk measure, this study aims to investigate whether the best estimation models for *VaR* also will be consistent with the best estimation models for *ES*. In order to answer this research question, the focus of this study is broken down into three main goals:

First, find the best estimation models to assess *ES* through a backtesting approach.

Second, find the best models to forecast *VaR* through Kupiec test. Even though many similar researches have been conducted, the results are still inconclusive.

Third, compare the best models for *VaR* and *ES*, then pinpoint their similarities and differences.

1.3 Research Limitations

The first limitation for this research derives from the limited empirical research as well as literature on the subject of backtesting *ES*. During the course of undergoing this research, there are only three viable backtesting approaches for *ES*. Amongst these three, one is considered to be the most applicable for this study. Given the role of backtesting in defining the best estimation models for *ES*, different choices of backtesting approach could lead to different conclusion of the best *ES* estimation models.

The second limitation would lie on the scope of the research. Given the time constraint to finish the research, the authors only base this study on the US market data, specifically the S&P500 index data.

The third limitation would be type of model chosen for this study. Given the time and technical constraints, some methods such as Age Weighted Historical Simulation (AWHS) and Volatility Weighted Historical Simulation (VWHS) are not included. However, the time-variation volatility aspect will still be captured by the other parametric methods.

The fourth limitation falls on the approach behind the conditional models in this study that is used to capture volatility clustering. For simplicity, this study only applies Exponentially Weighted Moving Average (EWMA) to capture volatility clustering. However, this might not be the most suitable approach to capture the actual behaviour of market return.

1.4 Outline of the Thesis

In general, this paper is divided into five main chapters. The first three chapters focus on the theoretical aspects of the thesis topic, while the last two chapters feature the empirical results of the study.

Chapter one highlights the fundamental concept behind the research. It introduces the research topic by stating the background information, underlying hypothesis and limitations, as well as defining the research problem along with its purpose and possible contribution.

Chapter two catalogues corresponding journals and literatures that relate to the research. This chapter present a brief but thorough investigation on the topic, which includes theoretical as well as empirical arguments that were made by the authors, fellow academicians and previous researchers.

Chapter three describes the approaches taken to conduct the research. It provides a detailed breakdown of the research methodologies and the data collection methods that will be used to accomplish the research purpose.

Chapter four elaborates on the findings obtained from the final results of the research. Based on the three goals of the research, the analysis of the findings are organised into three sections. Chapter four will start off with a discussion on the best models for *ES*, then continue with a dissection on the best models for *VaR*, and close off with a comparison of the best models.

Chapter five seals off the research by summing up the entire research outcome and answering the research problems. It also highlights the areas of improvement that exist in the study, which could be developed in future research.

2 Literature/Theoretical Review

In the field of risk management, *VaR* and *ES* are the two most commonly employed measures of risk. However, due to the later introduction of *ES*, the knowledge that surrounds *ES* is not as comprehensive as *VaR*, especially in terms of empirical research, model comparison and backtesting methodology.

Initially, financial institutions used *VaR* to control internal risk and to manage assets. Throughout its period as a standard risk measure, much empirical research has been conducted to determine the best models for the *VaR* approach. Unfortunately, the debates on which models would be best for *VaR* are still inconclusive.

[Pritsker \(2006\)](#) indicated that the Historical Simulation is not a good model to appraise *VaR* since this method adopts conservative estimates. [Pattarathammas et al. \(2008\)](#) explained that even though this non-parametric approach does not impose model risk and is simple to implement, HS possess several disadvantages. [Pattarathammas et al. \(2008\)](#) stated, “[In HS] an extrapolation beyond past observations tends to be inefficient with high variance estimator especially for small observations. Moreover, to mitigate the above problems by considering longer sample periods, the method tends to face the time-varying of return volatility”.

Alternatively, [Mandelbrot \(1963\)](#) and [Fama \(1965\)](#) rejected the Normal distribution model because the distribution of stock returns has fat tails. [McNeil \(1997\)](#), [Jondeau and Rockinger \(2003\)](#) and [Da Silva and Mendez \(2003\)](#) reported that during an extreme event, returns do not follow the Normal distribution because its empirical distribution has heavier tails. Therefore, classical parametric approach that is based on the assumption of Normal distribution is not suitable to estimate *VaR* during an extreme event like major financial crisis. Instead, [Blattberg and Gonedes \(1974\)](#) suggested to adopt Student t-distribution model because it has the appeal of fat tails. Despite of being a fat-tailed model, the Student t-distribution will not be the right model of estimation during extreme period. According to [Rachev & Mittnik \(2000\)](#), the Stable Paretian model, another fat-tailed model, would be a

better model during extreme period because its tails decay more slowly than the tails of the Normal distribution, which could better capture the extreme events that present in the data.

[Bollerslev \(1986\)](#) proposed a conditional heteroskedasticity model, the GARCH model. From that on, the dynamic methods have gained much popularity, since they can capture the stylized facts, volatility clustering and leptokurtosis, within the financial data. However, according to a research conducted by [Danielsson and de Vries \(2000\)](#), financial institutions preferred unconditional models due to their simplicity, although these models are based on false assumption of independence and equal distribution of returns. [Gilli and K llezi \(2006\)](#) admitted that the choice between conditional and unconditional model should depend on the period for the analysis and risk measures, of which the researcher or manager wants to use. [Baran and Witzany \(2011\)](#) compared Extreme Value Theory (EVT) methods to standard methods (Variance-Covariance and Historical Simulation) to calculate VaR and ES and showed that EVT methods, due to their nature, produced reasonable ES/VaR ratios. [Harmantiz et al. \(2006\)](#) and [Marinelli et al. \(2007\)](#) compared EVT based model with the Gaussian models, HS models, and Stable Paretian model, another heavy-tail approach. Their studies showed that models that could capture fat-tail have higher accuracy than those that do not. Moreover, they also proved that EVT models, especially the ones that adopted GPD method, raises more benefit. [Echaust and Just \(2013\)](#) found that conditional models perform better than unconditional models, and, the GARCH-EVT model enables to estimate the VaR correctly regardless of the considered assets.

As the significance of VaR as a risk measure grows, more problems of VaR are exposed. [Artzner et al. \(1997, 1999\)](#) cited that VaR is unable to capture “tail risk” and is not coherent, especially not subadditive. Subadditivity refers to the diversification effect in a portfolio. A risk measure that failed to fulfil the axiom of subadditivity property does not encourage diversification. Axioms are used to capture the intrinsic nature of a concept in a minimal yet precise manner. Hence, it is often used to depict a complex concept into a mathematical formulation. [Acerby and Tasche \(2001\)](#) emphasized the importance of having a coherent risk measure:

The axioms of coherence simply embody in a synthetic and essential way those features that single out a risk measure in the class of statistics of a portfolio dynamics.
... Broken axioms always lead to paradoxical, wrong results. (p.2)

Furthermore, [Yamai and Yoshihara \(2005\)](#) illustrated how *VaR*'s inability to capture tail risk can cause serious problem, as it may mislead rational investors who want to maximize their expected utility. Other researcher, [Liang and Park \(2007\)](#), confirmed [Artzner et al. \(1999\)](#) argument by providing empirical evidence which showed that *ES* is a superior downside risk measure compared to *VaR*. [Kerkhof and Melenberg \(2004\)](#) stated that a simulation study has provided evidence that “tests for expected shortfall with acceptable low levels have a better performance than tests for value-at-risk in realistic financial sample sizes.” (p.1). According to [Oh and Moon \(2006\)](#), *ES* estimates will always have bigger values than *VaR* estimates for the same critical level, which indicate that *VaR* has the tendency to underestimate tail-related risks. [Hürlimann \(2008\)](#) claimed that *ES* should replace *VaR* for the qualification of the reinsurance counterparty default risk in the Solvency II project, because *VaR* might lead to wrong comparative statistics results.

The Basel Committee on Banking Supervision (2014) also supported this argument by changing the standard risk measure from *VaR* to *ES* and stated that *ES* could assess tail risk “in a more comprehensive manner, considering both the size and the likelihood of losses above a certain threshold” (p.18). Though *ES* could overcome the problems with *VaR*, it also brings forth other issues. [Yamai and Yoshihara \(2002\)](#) mentioned that *ES* requires larger samples than *VaR* for the same level of accuracy. Moreover, [Gneiting and Ranjan \(2011\)](#) proved that *ES* is not elicitable as opposed to *VaR* through comparing density forecasts of *VaR* and *ES* based on threshold and quantile-weighted versions. The discovery of non-elicitability of *ES* led many to conclude that *ES* cannot be used to rank different point forecasts in a decision theoretically way, i.e., *ES* would not be backtestable ([Ziegel, 2014](#)). This misconception has become the last obstacle for the Basel Committee to fully implement *ES* as the standard regulatory risk measure.

Fortunately, many researchers have started to propose backtesting approaches for *ES*, such as [McNeil and Frey \(2000\)](#), [Berkowitz \(2001\)](#), [Wong \(2008\)](#). Researchers [Righi and Cerettea \(2015\)](#) applied the [McNeil and Frey \(2000\)](#) test to evaluate the accuracy of *ES* forecasting models, and then found that EVT models submitted incorrect quantile estimation, while unconditional fat-tailed models, the Filtered Historical Simulation models (FHS) and Indirect GARCH Conditional Autogressive Regression based model (IG-CARES) presented large p-values for the adopted tests. Meanwhile, [Righi and Cerettea \(2015\)](#) also observed that *VaR* estimation is crucial for the *ES* estimation, because the incorrect violation rate for *VaR*

estimates accompany with low p-values for the *ES* backtests. [Acerbi and Szekely \(2014\)](#) proved that elicibility is relevant for model selection, but irrelevant for model testing. Consequently, it is irrelevant for the choice of a regulatory risk standard. This duo also introduced three non-parametric methodologies to backtest *ES* that are easy to implement and would display better power than the standard Basel *VaR* test. These methodologies are free from distributional assumptions other than continuity, which is a necessary condition for any applications in banking regulations ([Acerby & Szekely, 2014](#)).

Some previous research on estimating *ES* with parametric approach has been conducted. However, due to the nature of the parametric approach that requires certain assumption on distribution or model, which might not accurately depict the actual behaviour of losses, the results could be unreliable. [Jadhav et al. \(2009\)](#) claimed that the use of parametric approach would lead to either underestimation or overestimation of *ES* because the approach relies on certain assumption of the distribution, which might not reflect the actual behaviour of the data (or in other word biased). [McNeil and Frey \(2000\)](#) along with [Harmantzis et al. \(2006\)](#) proposed and used the EVT approach to estimate *ES*. These researchers adopted a parametric inference within the framework of Generalized Pareto distribution (GPD) to estimate *ES*. Unfortunately, the implementation of this method was not smooth due to the scarcity of extreme data and problem with choosing a suitable method to estimate the parameters. Moreover, [Jadhav et al. \(2009\)](#) pointed out that a number of researches have showed that Gaussian approach tends to be unreliable in terms of estimating *ES*, because it does not seem to follow the normality assumption; it generates higher probability of large losses or gains compared to the probability implied by the normal assumption.

Previous research, such as [Acerbi and Tasche \(2002\)](#), [Koji and Kajima \(2003\)](#), [Harmantzis et al. \(2006\)](#) and [Song \(2008\)](#), have labelled Historical Simulation as the best method to estimate *ES*. [Harmantzis et al. \(2006\)](#) compared Historical Simulation approach with all other parametric methods available to estimate *ES* and found that HS approach successfully excels all the other approaches. On the other hand, [Song \(2008\)](#) conducted research that compared HS approach with another non-parametric approach called Kernel Based Method, a method that smoothens the data used to estimate *ES*. According to her research, [Song \(2008\)](#) found that the mean square error of the HS estimators is much smaller than the Kernel estimator; ergo, HS is the better estimator than the Kernel Method. [Odening and Hinrichs \(2003\)](#) as well as [Jadhav et al. \(2009\)](#) cited that HS approach is free of models and explicit assumptions of

return distributions because it uses the empirical distribution, which is generated from historical return data. Therefore, this model is considered to be robust, clear and easy to implement. However, HS is unduly sensitive to outliers. This will lead to an over-estimation of *ES* in the presence of outlier in the data.

Several authors are interested in the comparison between the best forecasting *VaR* models and the best estimating *ES* models. [Žiković and Dizdarević \(2011\)](#) conducted a research on the compatibility and interaction between *VaR* and *ES* as a risk measure for the fossil fuels prices within the energy market. Nine models are used to estimate *VaR* and *ES* from the two years daily return data on energy commodities. The estimation models for *VaR* are Normal Variance-Covariance Approach (VCV), RiskMetrics model, GARCH model, Historical Simulation (HS), Mirrored Historical Simulation (MHS), time weighted Boudoukh Richardson and Whitelaw Historical Simulation (BRW), Filtered Historical Simulation (FHS), Extreme Value approach using Generalized Pareto distribution (GPD) and conditional Extreme Value approach. Moreover, the nine estimation models for *ES* are VCV, RiskMetrics and GARCH with approach Gumbel distribution, bootstrapped Historical Simulation, bootstrapped Mirrored Historical Simulation, bootstrapped BRW and FHS approach, unconditional GPD and conditional Extreme Value (EVT-GARCH) approach. They employ the Kupiec test and the Christoffersen independence test to backtest *VaR*, and for *ES*, they only compare the performance of models but do not employ any backtesting methods. The empirical result of this study showed that the best performing *VaR* models are almost identical to the best performing *ES* models.

Then, [Aloui and Hamida \(2014\)](#) also administered a resembling research on forecasting *VaR* and *ES* for short and long trading positions within the Gulf Cooperation Council (GCC) stock markets. This study employs two Long Memory (LM) GARCH-class models under three alternative distributions: Normal, Student t and skewed Student t. They apply the Kupiec test and Dynamic Quantile test (DQT) to measure the accuracy of *VaR* estimates, and they connect the best model to forecast *ES* with the smallest mean squared error that the squared difference of the losses using the *ES* given that the *VaR* does not provide any information regarding the size of the expected loss. These researchers found that taking the long-memory in volatility, the asymmetry and the heavy-tails in stock returns into account is proven to enhance the performances of *VaR* and *ES*. They also showed that the LM ARCH/GARCH-class models are good choices to estimate *VaR* and *ES*.

Similar to previous research, this study also aims to distinguish the accuracy and compare the effectiveness of *VaR* and *ES* as risk measure. However, it focuses on different market with longer observation periods, which are deliberately chosen to apprehend the impact of the different economic situations on the model estimations. This study uses nine models to appraise *VaR* and *ES*. The major difference between this study and the previous study is the adaptation of ES backtesting.

3 Methodology

Value-at-risk is the smallest loss ℓ such that the probability of a future portfolio loss L that is larger than the loss ℓ , is less than or equal to $1 - \alpha$. Mathematically define by:

$$VaR_\alpha(L) = \min\{\ell: Pr(L > \ell) \leq 1 - \alpha\}, \quad (1)$$

where α is the tolerance level or confidence level of VaR .

[Artzner et al. \(1997\)](#) proposed expected shortfall to solve the inherent problems existing in VaR . ES at confidence level α is the conditional expectation of loss given that the loss is larger than VaR_α , which can be statistically defined as the average VaR for confidence levels larger than or equal to α :

$$ES_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_x(L) dx. \quad (2)$$

Basel chooses $VaR_{0,99}$ and $ES_{0,975}$.

3.1 Estimation Approach

3.1.1 Historical Simulation (HS)

Contributed by its simplicity, transparency, a nature of being free of distributional assumption, able to capture fat tails and able to produce accurate forecasts, HS approach has become popular. [Perignon and Smith \(2010\)](#) found that 73% of the banks that disclose their value at risk on their 2005 annual reports use HS or its variants.

The calculation of VaR and ES under Historical Simulation directly relies on the sample of observed losses. This means that the VaR estimates are calculated directly from the values of the empirical profit and loss distribution. Therefore, it does not require any explicit assumptions on the return distribution ([Odening and Hinrichs, 2003](#)).

For a sample of N losses, approximately $(1 - \alpha)N$ losses that are larger than value-at-risk are expected to be found. So the $(1 - \alpha)N + 1$ largest loss in the sample is the estimate of VaR_α , assuming that $(1 - \alpha)N$ is an integer. According to the definition of ES , $ES_\alpha(L)$ equals to the average of the $(1 - \alpha)N$ largest losses.

Given the time and technical constraints, only this unconditional non-parametric method will be used, other non-parametric methods such as AWHs and VWHs will not be included. However, the time-variation volatility aspect will still be captured by the other parametric methods.

3.1.2 Gaussian Distribution (N-dist)

The Gaussian distribution is characterized by mean (μ) and variance (σ^2). Sometimes, financial intuitions just assume the losses follow a Normal distribution and use historical data to estimate the model parameters (μ and σ^2). The probability density function is of the form:

$$f_{norm}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, \quad (3)$$

$$f_{std}(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}. \quad (4)$$

According to the definition of VaR , VaR and ES can be estimated using the following formulas:

$$VaR_\alpha(L) = \mu + \sigma z_\alpha, \quad (5)$$

$$ES_\alpha(L) = \mu + \sigma \frac{f_{std}(z_\alpha)}{1-\alpha}, \quad (6)$$

where z_α is the α quantile of the standard Normal distribution.

Under the assumption of Normal distribution, VaR provides the same information on tail loss as does ES in the sense that both are proportional to volatility, since z_α and $\frac{f_{std}(z_\alpha)}{1-\alpha}$ are constant once α is fixed. For example, given that $\alpha = 0,99$ and $\mu = 0$,

$$VaR_{0,99}(L) = 2,326\sigma, \quad (7)$$

$$ES_{0,99}(L) = 2,67\sigma. \quad (8)$$

The conditional model is the one allowing for a time varying volatility. So, VaR and ES under Normal distribution are:

$$VaR_{\alpha}(L) = \mu + \sigma_{T+1} z_{\alpha} , \quad (9)$$

$$ES_{\alpha}(L) = \mu + \sigma_{T+1} \frac{f_{std}(z_{\alpha})}{1-\alpha} , \quad (10)$$

where μ is still the sample mean, but σ_{T+1} is the forecasted volatility for the next period using the GARCH or the EWMA model.

3.1.3 Student t-Distribution (t-dist)

It is a well-established fact that financial returns have fat, heavy or long tails, which means that kurtosis (k) is larger than for the Normal distribution, which is always equal to exactly 3. If kurtosis is larger than 3, VaR under Normal distribution underestimates risk. The reason for introducing the Student t-distribution as an alternative relies on its ability accommodate kurtosis that is larger than 3. Kurtosis can be captured by an additional parameter ν (degree of freedom). For $\nu > 4$,

$$\nu = \frac{4k-6}{k-3} . \quad (11)$$

The probability density function for a Student t-distribution is of the form:

$$f(x) = \frac{\Gamma[(\nu+1)/2]}{\sigma\sqrt{(\nu-2)\pi}\Gamma(\nu/2)} \left[1 + \frac{1}{\nu-2} \left(\frac{x-\mu}{\sigma} \right)^2 \right]^{-(\nu+1)/2} , \quad (12)$$

where μ is the mean and σ is the standard deviation.

Then, the estimates of VaR and ES under Student t-distribution are obtained as:

$$VaR_{\alpha}(L) = \mu + \sqrt{\frac{\nu-2}{\nu}} \sigma t_{\alpha,\nu} , \quad (13)$$

$$ES_{\alpha}(L) = \mu + \sqrt{\frac{\nu-2}{\nu}} \sigma \frac{f_{std}^*(t_{\alpha,\nu})}{1-\alpha} \left(\frac{\nu+t_{\alpha,\nu}^2}{\nu-1} \right) , \quad (14)$$

where f_{std}^* is the probability density function for a standardized t-distributed variable, i.e. with

$\mu = 0$ and $\sigma^* = 1$ and $t_{\alpha,v}$ is the α quantile for this distribution. σ^* is a parameter closely related to standard deviation:

$$\sigma^* = \sqrt{\frac{v-2}{v}} \sigma. \quad (15)$$

VaR and ES under the conditional Student t-distribution are:

$$VaR_\alpha(L) = \mu + \sqrt{\frac{v-2}{v}} \sigma_{T+1} t_{\alpha,v}, \quad (16)$$

$$ES_\alpha(L) = \mu + \sqrt{\frac{v-2}{v}} \sigma_{T+1} \frac{f_{std}^*(t_{\alpha,v})}{1-\alpha} \left(\frac{v+t_{\alpha,v}^2}{v-1} \right), \quad (17)$$

where σ_{T+1} is the forecasted volatility for the next period.

3.1.4 Extreme Value Theory (EVT)

The Generalized Pareto distribution (GPD) is the key distributions of Extreme Value Theory (EVT). The idea underlying EVT is to model the extreme outcomes rather than all outcomes because it is exactly these large losses that are relevant for estimating VaR (and ES).

The Peak Over Threshold model (POT) has become the preferred extreme value approach in finance. By using all losses in a sample larger than some pre-specified threshold value, POT solves the problem of information loss that happens in traditional EVT.

The theory underlying the POT approach aims to model the excess losses $L - u$, where u is the predetermined threshold value. Let F be the cumulative density function of the stochastic loss variable L , i.e. $\Pr(L \leq \ell) = F(\ell)$. Then define a cumulative density function $F_u(\ell)$ for excess losses $L - u$ given that L exceeds u :

$$F_u(\ell) = \Pr(L - u \leq \ell | L > u) = \Pr(L \leq \ell + u | L > u) = \frac{F(\ell+u) - F(u)}{1 - F(u)}. \quad (18)$$

Then,

$$F_u(\ell - u) = \frac{F(\ell) - F(u)}{1 - F(u)}. \quad (19)$$

Solving for $F(\ell)$,

$$F(\ell) = [1 - F(u)]F_u(\ell - u) + F(u). \quad (20)$$

To solve for the α quantile VaR_α , Pickands-Balkema-deHaan extreme value theorem ([Pickands, 1975](#); [Balkema & de Haan, 1974](#)) says that the distribution function $F_u(\ell - u)$ converges to a Generalized Pareto distribution $G(\ell - u)$ provided that u is sufficiently large.

The probability density function for a GPD can be written as:

$$G(\ell - u) = \begin{cases} 1 - \left(1 + \xi \frac{\ell - u}{\beta}\right)^{-\frac{1}{\xi}}, & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\frac{\ell - u}{\beta}\right), & \text{if } \xi = 0 \end{cases}, \quad (21)$$

where ξ is a shape parameter called the tail index in the Generalized Pareto distribution. A higher value of ξ means that there is more probability in the right tail.

Using the approximation of $F_u(\ell - u) \approx G(\ell - u)$, putting $F(VaR_\alpha) = \alpha$, and estimating $F(u)$ with $(N - N_u)/N$, the POT estimate of VaR and ES become:

$$VaR_\alpha = \begin{cases} u + \frac{\beta}{\xi} \left[\left(\frac{N}{N_u} (1 - \alpha) \right)^{-\xi} - 1 \right], & \text{when } \xi \neq 0 \\ u - \beta \ln \left(\frac{N}{N_u} (1 - \alpha) \right), & \text{when } \xi = 0 \end{cases}. \quad (22)$$

According to the definition of ES_α :

$$ES_\alpha = \frac{1}{1 - \alpha} \int_\alpha^1 VaR_x dx. \quad (23)$$

Then: (see [Appendix A](#))

$$ES_\alpha = \frac{\beta - u\xi}{1 - \xi} + \frac{VaR_\alpha}{1 - \xi}, \quad (24)$$

where β is a scale parameter and both parameters (β and ξ) can be estimated by ML.

The limit theorem makes no specific statement about the underlying (parent) distribution for the losses; a GPD distribution is 'always' a proxy of the limiting distribution of $F_u(\ell - u)$.

EVT also can be used in a conditional model through combing POT model with GARCH/EWMA volatility models. The dynamic approach models the conditional loss distribution to forecast the loss over the next period. [McNeil and Frey \(2000\)](#) proposed a dynamic VaR forecasting method based on EVT, in which the GARCH method is used to model the current market volatility background.

Still assume that ℓ be the loss variable:

$$\ell_t = \bar{\ell} + \sigma_t \varepsilon_t^*, \quad t = 1, 2, \dots, T, \quad (25)$$

where $\bar{\ell}$ is the sample average of the observed losses, σ_t is volatility at time t , and ε_t^* denotes the standardized residual.

According to the ‘‘hybrid method’’, VaR and ES under conditional POT can be estimated using the following formulas:

$$VaR_\alpha = \mu + \sigma_{T+1} VaR(\varepsilon^*), \quad (26)$$

$$ES_\alpha = \mu + \sigma_{T+1} ES(\varepsilon^*). \quad (27)$$

$VaR(\varepsilon^*)$ and $ES(\varepsilon^*)$ are the values at risk and expected shortfall for the standardized residuals under the GPD distribution at confidence level α , respectively. σ_{T+1} denotes the GARCH/EWMA volatility forecast one period out of sample.

3.2 Model Evaluation

The accuracy of the results obtained using the proposed models should be verified via backtesting. Regulators are principally concerned about underestimated VaR or ES , which will cause a problem because then banks will report to low risk as well as their regulatory assets reserved. Therefore, to make sure that the models do not yield underestimated results, the one-sided test is used. The null hypothesis of this one sided test is that the model is correct; while, the alternative hypothesis is that the VaR or ES is underestimated. The explanations on how to use backtesting methods to verify the accuracy of different models are given in the following sections.

3.2.1 Backtesting *VaR*: Kupiec Frequency Test

[Kupiec \(1995\)](#) considered statistical techniques that can be used to quantify the accuracy of the tail values of the distribution of losses. The Kupiec frequency test is the most fundamental test, stating that the actual number of *VaR* violations, call it x , is significantly different from the expected number of *VaR* violations, which is $(1 - \alpha)N$, where N is the total number of observations.

A *VaR* violation is said to occur if the observed loss exceeds our *VaR* estimate for a given day. The actual number of violations (x) is binomially distributed:

$$\Pr(X = x) = \binom{N}{x} p^x (1 - p)^{N-x}, \quad (28)$$

where $p = 1 - \alpha$.

The cumulative probabilities are then:

$$\Pr(X \leq x) = \sum_{i=0}^x \binom{N}{i} p^i (1 - p)^{N-i}. \quad (29)$$

If the actual frequency of violations deviates too much from the predicted frequency of violations, the model underlying the *VaR* estimator is statistically rejected.

The one-sided Kupiec test is used to check if the actual frequency of violations is “too large” compared to the expected frequency of violations, because regulator only care about the risk underestimation problem. As a result, calculate $\Pr(X \geq x) = 1 - \Pr(X < x) = 1 - \Pr(X \leq x - 1)$ no matter the actual number of violations is larger than or smaller than the expected number of violations, and then compare this probability to the statistical level of interest, e.g. 5%. If $\Pr(X \geq x) <$ the statistical level of interest, the model is statistically rejected.

3.2.2 Backtesting ES

[Acerbi and Szekely \(2015\)](#) proposed three backtesting methods for expected shortfall. Amongst these three, the second method—test statistic Z_2 , which is referred to as “Testing ES Directly”, is considered to be the most applicable one because the other two tests require Monte Carlo simulation of the distribution of the test statistic to compute the p-value and therefore need to store predictive distributions. Moreover, based on fixed significance thresholds; Z_2 can be treated as a traffic-light system, in which it shows a remarkable stability of the significance thresholds across a wide range of tail index values, which spans over all financially realistic cases. Besides, calculating Z_2 only requires recording two numbers per day: one is the estimated $ES_{\alpha,t}$ and the other one is the magnitude $L_t I_t$ of a $VaR_{\alpha,t}$ exception, where L_t is the loss at time t and I_t is an indicator variable:

$$I_t = \begin{cases} 1, & \text{when } L_t > VaR_{\alpha,t} \\ 0, & \text{otherwise} \end{cases} \quad (30)$$

In reality, it is sufficient to only record the size of the α – tail of the model distribution, since $L_t I_t$ can be simulated because of $I_t \sim Bernoulli$.

Similar to the previous notations, the independent profit loss is denoted as L_t , the true but unknown distribution of profit losses is F_t , and the model distribution is represented by P_t . Assume that the losses follow a continuous distribution. Then, the definition of ES can be written as:

$$ES_{\alpha,t} = \mathbb{E}[L_t | L_t > VaR_{\alpha,t}], \quad (31)$$

where α is the confidence level and the Basel governed $\alpha = 0,975$.

We may rewrite the formula assuming that the model is correct:

$$ES_{\alpha,t} = \mathbb{E} \left[\frac{L_t I_t}{1-\alpha} \right]. \quad (32)$$

Now, define the test statistic:

$$Z_2(L) = - \sum_{t=1}^T \frac{L_t I_t}{T(1-\alpha)ES_{\alpha,t}} + 1. \quad (33)$$

The null hypothesis states that the model exact in the tail, while the alternative hypothesis states that ES is underestimated. They are mathematically written as:

$$H_0: P_t^{[\alpha]} = F_t^{[\alpha]} \text{ for all } t$$

$$H_1: ES_{\alpha,t}^P \leq ES_{\alpha,t}^F \text{ for all } t \text{ and } < \text{ for some } t$$

$$VaR_{\alpha,t}^P \leq VaR_{\alpha,t}^F \text{ for all } t$$

Under the null hypothesis,

$$\begin{aligned} \mathbb{E}_{H_0}[Z_2(L)] &= \mathbb{E}_{H_0}\left[-\frac{1}{T}\sum_{t=1}^T \frac{L_t I_t}{(1-\alpha)ES_{\alpha,t}} + 1\right] \\ &= -\frac{1}{T}\sum_{t=1}^T \mathbb{E}\left[\frac{L_t I_t}{1-\alpha}\right] \frac{1}{ES_{\alpha,t}} + 1 \\ &= -\frac{1}{T}\sum_{t=1}^T ES_{\alpha,t} \frac{1}{ES_{\alpha,t}} + 1 \\ &= -\frac{1}{T}\sum_{t=1}^T 1 + 1 \\ &= 0 \end{aligned} \tag{34}$$

Then, we can conclude:

$$\mathbb{E}_{H_0}[Z_2] = 0 \quad \text{and} \quad \mathbb{E}_{H_1}[Z_2] < 0$$

Provided the critical values for different significance levels and different degree of freedom (see *Table 1*), which was published by [Acerbi and Szekely \(2015\)](#), it is clear that the thresholds deviate significantly from $-0,7$ only for dramatically heavy tailed distribution, with ν closing to 3. As a result, Z_2 with fixed levels $Z_2^* = -0,7$ (when significance level is 5%) and $Z_2^* = -1,8$ (when significance level is 0,01%) would perfectly be the traffic-light in all occasions, which implies that Z_2 lends itself to implementations that do not require the recording of the predictive distributions. The ± 1 location shifts across an unrealistically large area for a real loss distribution, which is expected to converge around zero.

Finally, compare the actual value of Z_2 to the critical value. If Z_2 is smaller than the corresponding critical value, then the underlying model is rejected.

Table 1: 5% and 0,01% significance thresholds for Z_2

v	Significance					
	5%			0,01%		
	Location			Location		
	-1	0	1	-1	0	1
3	-0,78	-0,82	-0,88	-3,9	-4,4	-5,5
5	-0,72	-0,74	-0,78	-1,9	-2,0	-2,3
10	-0,70	-0,71	-0,74	-1,8	-1,9	-1,9
100	-0,70	-0,70	-0,72	-1,8	-1,8	-1,9
Gaussian	-0,70	-0,70	-0,72	-1,8	-1,8	-1,9

Source: Acerbi and Szekely (2015)

3.3 Data Analysis

The models are tested using daily data of the S&P500 index, collected from Yahoo! Finance, from January 1, 2005 to December 31, 2014. Then, the continuously compounded returns are calculated. The returns are calculated as 100 times the first difference of the natural logarithm of each of the series, $R_t = 100\ln(p_t/p_{t-1})$, where R_t denotes the return of index for date t and p_t represents the index price at time t . Finally, the loss $L_t = -R_t$ is defined. *Figure 1* presents the loss observations from 2005 to 2014. There is a huge fall in the market from the middle of 2008 to 2009, which increases the market risk of the portfolio.

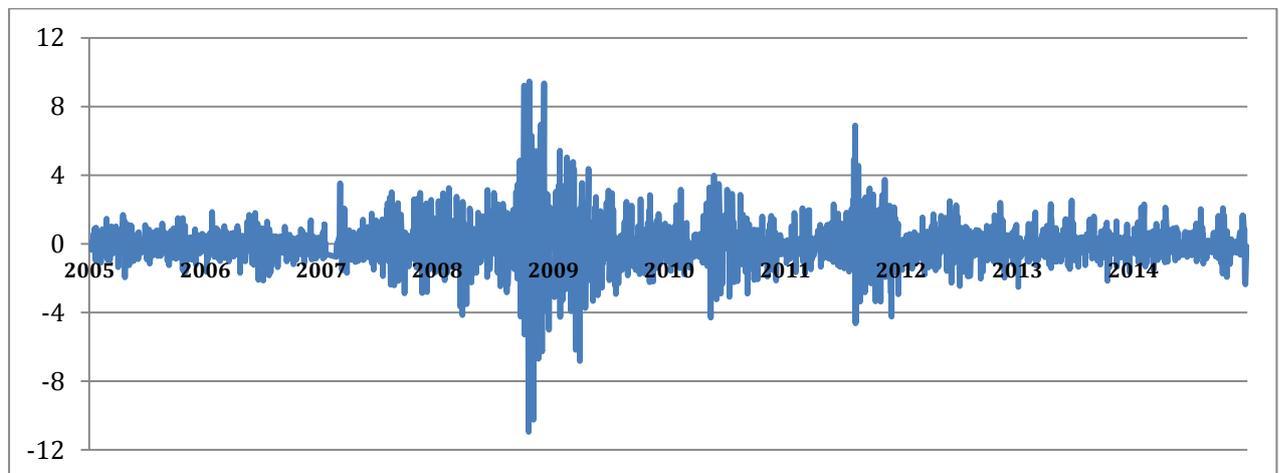


Figure 1: The loss observations for the entire sample period

The first estimation window is from 2005 to 2006. Then, the test period is from 2007 to 2014. In order to assess the financial crisis effect, the evaluation period is divided into three parts: Bad Economy Period (2007 - 2008), Recovery Period (2009 - 2011), and Good Economy Period (2012 - 2014). *Table 2* presents the summary statistics for losses over the entire sample period 2005-2014, which includes the first estimation window (2005-2006) and the three testing periods. There are a total of 2514 observations obtained from the entire sample. For all the considered models, rolling windows with the size of 500 days are used to calculate new value of *VaR* or *ES* as a risk forecast for the next trading day.

Table 2: Loss observations: Descriptive Statistics

	2005 - 2006	Testing Periods		
		2007 - 2008	2009 - 2011	2012 - 2014
Mean	-0,03616	0,08953	-0,04378	-0,06538
Standard Error	0,02850	0,08754	0,05310	0,02695
Median	-0,08070	-0,06056	-0,11070	-0,06609
Standard Deviation	0,63735	1,96521	1,46009	0,74013
Sample Variance	0,40621	3,86207	2,13187	0,54780
Kurtosis	0,51823	7,23108	2,73603	1,18237
Skewness	-0,03510	0,18716	0,23242	0,22259
Range	3,98321	20,42671	13,73247	5,04144
Minimum	-2,13358	-10,95720	-6,83664	-2,50860
Maximum	1,84963	9,46951	6,89583	2,53284
Sum	-18,07801	45,12149	-33,09611	-49,29667
Observations	500	504	756	754

4 Analysis and Discussion

This chapter reports the *ES* and *VaR* forecasts based on the models described in chapter 3. It also contains the details of the examination and comparison of the best forecasting models.

4.1 *ES* Estimates and Backtesting

To implement conditional *ES* and *VaR* estimation, EWMA with $\lambda = 0,94$ is used to capture the volatilities, because across a range of different market variables EWMA with $\lambda = 0,94$ gives forecasts of the variance rate that come closest to the realised variance rate ([JPMorgan, 1996](#)).

For the EVT, Maximum Likelihood Method is used to obtain the estimates of parameters ξ and β . Parameters are estimated "before" every new evaluation period starts. Considering the fact that EVT only focus on the large losses and these large losses are essentially the same before and after the estimation window is moved one day or one month forward, the estimates of EVT parameters would be almost the same, and consequently there is no reason to update parameters too often. Since a value close to the 95th percentile point of the empirical distribution usually works well ([Hull, 2012](#)), the 95% point is chosen as the threshold value u . Observations are then sorted from the highest to the lowest value. Then, the observations above u are selected and the GPD is fitted to excess returns. The function Solver from Excel application is used to search for the values of ξ and β that maximize the corresponding log-likelihood function; the parameters estimates results are shown in [Appendix B](#).

Based on the presented models and the estimated parameters, series of forecasted *ES* at confidence level 97,5% are obtained. To emphasize the difference under the distinctive estimation models, temporal evolutions of the estimated *ES* for the S&P 500 under each model: unconditional HS, conditional and unconditional Normal distribution, conditional and unconditional Student t-distribution, together with conditional and unconditional EVT, are generated and observed (see [Appendix C](#)).

The estimated *ES* under the HS, unconditional Normal distribution and unconditional Student t-distribution models showed similar processes within the evaluation window: *ES* rises steadily during the bad economy period (2007-2008) and shows a drastic increase from September 2008 until the end of the year. This is caused by a huge fall in the market during this period, which led to higher risk levels. *ES* remains at its highest in 2009 to 2010, and then gradually decreases from 2011.

The *ES* determined by the unconditional EVT (including the situation of $\xi = 0$ and $\xi \neq 0$) show different evolution process: the forecasted *ES* seems to be at its highest during the bad economy period, which might result from using only extreme high losses found in 2008 to estimate *ES*. The *ES* for the conditional models (N-dist, t-dist and EVT) overall shows similar fluctuation patterns with the unconditional models (HS, N-dist, t-dist).

After completing the *ES* estimations, the backtesting approach is conducted. *Table 3* highlights the result of the *ES* backtesting at 5% significance level, the same significance level as used in the Basel traffic-light mechanism. As mentioned before, if a 5% significance level is used, the traffic light is based on Z_2 with fixed level $-0,7$, which can perfectly be applied in all situations. All negative Z_2 can be seen as the result of *ES* underestimation. Since a value lower than $-0,7$ will be rejected, then the negative Z_2 can be divided into two parts: the confidence interval $[-0,7, 0]$, and the underestimated within the rejection region $[-\infty, -0,7]$. Only if the Z_2 is falling in the rejection region the underestimation is statistically significant. It is very important to note that *ES* backtesting is a one-sided test that focuses on the underestimation of risk. Therefore, overestimation of risk, represented by positive Z_2 , does not lead to a rejection of the null hypothesis.

First, the performance of the models is examined thoroughly under each period. During the bad economy period, *ES* is significantly underestimated by most of the models such as the HS, unconditional and conditional Normal distribution models, also unconditional and conditional Student t-distribution models, and therefore these models are rejected, contradicting [Acerbi and Tasche \(2002\)](#), [Koji and Kajima \(2003\)](#), [Harmantzis et al. \(2006\)](#) and [Song \(2008\)](#). On the contrary, EVT models show relatively good results for *ES* estimation. The test statistics Z_2 falls in the confidence interval when *ES* estimates are determined by the unconditional EVT models. The conditional EVT model with $\xi = 0$ overestimates *ES* while the conditional EVT with $\xi \neq 0$ underestimates *ES*, however both models are accepted.

Moreover, during the recovery period, only the conditional Normal distribution model is rejected. During the good period, all conditional models (N-dist, t-dist and EVT) are rejected. Furthermore, in accordance to research conducted by [Danielsson and de Vries \(2000\)](#), it is not meaningful to make the models more dynamic by taking current market condition into account during the recovery and good economy period, since all unconditional models are not rejected during these two periods. This is due to the fact that unconditional volatilities have long memory of large losses that have happened in previous years, while the standard volatilities determined by EWMA capture the current market condition but give lesser weights to the previous large losses, and therefore underestimate *ES*.

Then, the realizations of the same model across different periods are elaborated. Firstly, even though the HS model is not rejected in recovery and good economy period, it overestimates *ES* during these periods. Meanwhile, it underestimates *ES* during the financial crisis, which is also known as the bad economy period. This drawback of HS was also well documented in [Odening and Hinrichs \(2003\)](#) and [Jadhav et al. \(2009\)](#).

Secondly, the magnitude of the *ES* estimates is larger when determined by the Student t-distribution models rather than by Normal distribution models, as corroborated by [Aloui and Hamida \(2014\)](#). Unfortunately, neither of these models works perfectly for all three periods. This suggests that neither Normal distribution nor Student t-distribution models are good choices when the period of the test window is uncertain. This result supports the finding of [Jadhav et al. \(2009\)](#).

Thirdly, EVT provides well-established statistical models for *ES*, in spite of the risk of overestimation in the recovery period, which corroborates [McNeil and Frey \(2000\)](#), [Marinelli et al. \(2007\)](#), [Jalal and Rockinger \(2008\)](#), [Baran and Witzany \(2011\)](#), [Wong et al. \(2012\)](#) and [Singh et al. \(2013\)](#), but contradicts [Tolikas \(2014\)](#) as well as [Righi and Ceretta \(2015\)](#).

Lastly, conditional EVT fails at the good economy period, while unconditional EVT successes in all periods, which implies that conditional EVT models are not necessary. The reason for this is that often EVT focuses on large losses, meaning that *ES* is estimated at a very high confidence level, with rarely corresponding tail-events, which only happens perhaps once every five or even every ten years.

Table 3: Backtesting Results for Expected Shortfall

		HS	N-dist	N-dist EWMA	t-dist	t-dist EWMA	EVT		EVT-cond	
							$\xi = 0$	$\xi \neq 0$	$\xi = 0$	$\xi \neq 0$
Bad Economy Period	Z_2	-3,522	-5,498	-1,769	-4,721	-1,352	-0,044	-0,052	0,200	-0,007
	Result of Test	reject	reject	reject	reject	reject	cannot reject	cannot reject	cannot reject	cannot reject
	Analysis	underestimate	underestimate	underestimate	underestimate	underestimate	underestimate	underestimate	overestimate	underestimate
Recovery Period	Z_2	0,304	-0,344	-1,057	-0,169	-0,644	0,007	0,038	0,096	0,054
	Result of Test	cannot reject	cannot reject	reject	cannot reject	cannot reject	cannot reject	cannot reject	cannot reject	cannot reject
	Analysis	overestimate	underestimate	underestimate	underestimate	underestimate	overestimate	overestimate	overestimate	overestimate
Good Economy Period	Z_2	0,457	0,151	-1,036	0,300	-0,773	-0,040	0,167	-1,785	-1,201
	Result of Test	cannot reject	cannot reject	reject	cannot reject	reject	cannot reject	cannot reject	reject	reject
	Analysis	overestimate	overestimate	underestimate	overestimate	underestimate	underestimate	overestimate	underestimate	underestimate

Note: The *ES* are forecasted at the confidence level of 97,5%, 5% Significance level and Critical value of $-0,7$. Negative Z_2 means underestimated *ES*, while positive Z_2 represents overestimation of risk.

4.2 VaR Estimates and Kupiec Tests

The estimated VaR under different models are presented in [Appendix D](#). The VaR that are determined by HS shares similar patterns with the VaR that are determined by unconditional Normal distribution model as well as by unconditional Student t-distribution model. These estimated VaR increases constantly from the beginning of 2007, then continues to peak at their highest level from 2009 up to 2010 because a bad day will increase the probability of large loss on the next day; ergo the VaR is expected to increase if the past returns have shown very negative values. After the crisis has recovered, the VaR seems to decline slowly.

The VaR values under unconditional EVT with $\xi = 0$ and $\xi \neq 0$ only vary slightly. The value of VaR under the model with $\xi = 0$ drops from 6,44 in the financial crisis (bad economy) period to 4,23 in the recovery period and finally settles at 1,97 in the good economy period. On the other hand, the value of VaR under the model with $\xi \neq 0$ drops from 6,50 to 4,26 and settles at 2,01. The VaR under conditional models have similar patterns, which are consistent with the pattern of volatilities captured by EWMA.

After a visual analysis of the similarities and differences among the forecasting models, Kupiec test is applied to these models to verify their accuracy. *Table 4* illustrates the results of the Kupiec test at 5% significance level. If the probability is larger than or equal to 5%, the model is accepted; otherwise, we reject this model from being the best VaR forecasting model. When the actual number of violations is larger than the expected number of violations, the VaR estimates are underestimated; otherwise, the VaR estimates are overestimated. Similar to ES backtesting, the one-sided Kupiec test that focuses on the underestimation of risk is employed. Therefore, overestimation of risk does not lead to a rejection of the null hypothesis. It is obvious that various estimation models and different testing periods could lead to varying results of VaR estimation.

During the bad economy period, all models except EVT are statistically rejected. This result is linked with the findings of [Gençay and Selçuk \(2004\)](#), [Harmantiz et al. \(2006\)](#) and [Marinelli et al. \(2007\)](#). These researchers found empirical evidence that VaR forecasts are more accurate if estimated by EVT-based models. The HS, unconditional and conditional Normal distribution as well as Student t-distribution models all significantly underestimate VaR and

therefore are concluded as having “too many” violations, which corroborates with [McNeil \(1997\)](#), [Jondeau and Rockinger \(2003\)](#) and [Da Silva and Mendez \(2003\)](#). During an extreme event, the empirical distribution of losses has heavier tails. Therefore, classical parametric approach that is based on the assumption of normal distribution or Student t-distribution is not suitable to estimate *VaR* during an extreme event like major financial crisis.

Nonetheless, when the evaluation window shifts to years 2009 and 2011, only the conditional Normal distribution model is rejected. After the financial crisis, HS maintains way too high estimates of *VaR* and consequently leads to risk overestimation, given that during this period, HS still utilises the large losses that occurred during the financial crisis period to calculate *VaR*. It is obvious that the unconditional models are accepted while all the other conditional models are rejected during the good economy period, which implies that the conditional models work worse for *VaR* estimation, contradicting to [Dowd \(2005\)](#) and [Echaust and Just \(2013\)](#).

After analysing the performance of models in each period, the accuracy of each model through the three periods is evaluated. Firstly, it is clear that the HS significantly underestimates *VaR* during the financial crisis because this method tends to maintain a very conservative level during bad economy period. On the contrary, HS overestimates *VaR* during the recovery and good economy periods. These findings imply that HS is not a good option to estimate *VaR*. This result is consistent with the discoveries of [Pritsker \(2006\)](#) and [Righi and Ceretta \(2015\)](#).

Secondly, Student t-distribution models always generate better results (i.e., fewer violations) than Normal distribution models in the same situation, as confirmed by [Al-Maghyreh and Awartani \(2012\)](#) and [Aloui and Hamida \(2014\)](#). The primary reason for this is that it is a well-established fact that the empirical distributions of financial losses have fatter, heavier or longer tails than the Normal distribution, and Student t-distribution can capture this fatter distribution. Accordingly, Normal distribution underestimates *VaR* compared to Student t-distribution. Nevertheless, none of them is good enough to forecast *VaR* in either period, since neither Normal distribution nor Student t-distribution pass the Kupiec test for every period. This result is consistent with [Mandelbrot \(1963\)](#) and [Fama \(1965\)](#), but against [Degiannakis \(2004\)](#).

Furthermore, the unconditional EVT model with $\xi = 0$ or $\xi \neq 0$ estimates *VaR* perfectly at all the time, as corroborated by [Baran and Witzany \(2011\)](#). More interestingly, the number of violations obtained under the normal EVT model is almost the same as that obtained under the corresponding EVT model with $\xi \neq 0$, and consequently refers to the same results.

Finally, the conditional EVT models are not significantly meaningful, considering that the unconditional model is enough to forecast *VaR*.

Table 4: Kupiec Backtest Results for Value-at-Risk

		HS	N-dist	N-dist EWMA	t-dist	t-dist EWMA	EVT		EVT-cond	
							$\xi = 0$	$\xi \neq 0$	$\xi = 0$	$\xi \neq 0$
Bad Economy Period	Nr. of violations	32	44	21	39	15	5	5	2	3
	Prob.	<0,001	<0,001	<0,001	<0,001	<0,001	0,391	0,391	0,880	0,742
	Result of Test	reject	reject	reject	reject	reject	cannot reject	cannot reject	cannot reject	cannot reject
	Analysis	underestimate	underestimate	underestimate	underestimate	underestimate	overestimate	overestimate	overestimate	overestimate
Recovery Period	Nr. of violations	5	11	18	6	11	10	10	8	8
	Prob.	0,766	0,082	<0,001	0,631	0,082	0,142	0,142	0,346	0,346
	Result of Test	cannot reject	cannot reject	reject	cannot reject	cannot reject	cannot reject	cannot reject	cannot reject	cannot reject
	Analysis	overestimate	underestimate	underestimate	overestimate	underestimate	underestimate	underestimate	underestimate	underestimate
Good Economy Period	Nr. of violations	2	9	21	5	17	10	10	30	30
	Prob.	0,981	0,227	<0,001	0,764	0,001	0,140	0,140	<0,001	<0,001
	Result of Test	cannot reject	cannot reject	reject	cannot reject	reject	cannot reject	cannot reject	reject	reject
	Analysis	overestimate	underestimate	underestimate	overestimate	underestimate	underestimate	underestimate	underestimate	underestimate

Note: *VaR* are estimated at the confidence level of 99%. The significance level used is 5%. The expected numbers of violations are 5,04, 7,56, 7,54 respectively for each period. When the actual number of violations is larger than the expected number of violations, the *VaR* estimates are underestimated; otherwise, the *VaR* estimates are overestimated.

4.3 Comparison of Best Models

According to the above analysis, the best performing *VaR* models are identical to the best performing *ES* models, just as the finding of [Žiković and Dizdarević \(2011\)](#). The unconditional EVT models are the only models that could successfully pass the backtesting process for all periods and for both *VaR* and *ES*.

Besides, taking a panoramic view of all the situations, the probability of risk underestimation for each model when adapted to estimate *ES* is smaller compared to that when used to forecast *VaR*. *Table 3* along with *Table 4* showed that 17 out of the 27 models have caused risk underestimation problem when used to forecast *ES*; meanwhile, 19 models lead to risk underestimation problem when used to estimate *VaR*. The underestimation of risk would lead to lower regulatory capital than needed.

The HS model underestimates both *VaR* and *ES* and is rejected during the financial crisis period. On the other hand, this method overestimates both risk measures, *VaR* and *ES*, and is accepted during the recovery periods.

Both unconditional Normal distribution model and unconditional Student t-distribution model fail in the backtesting process in the financial crisis (bad economy) period, but success in the other two periods. The result is still consistent, whether the models are used to estimate *VaR* or *ES*. Additionally, the dynamic parametric models perform worse than the corresponding unconditional models. Specifically, the conditional Normal distribution models, as it seems to underestimate *VaR* and *ES* and are also rejected in all periods. There are two key reasons contributed to this. First is that Normal distribution model does not consider the stylized fact of fat tail of financial data. Second is the fact that the standard EWMA volatilities capture the current market condition but give lesser weight to the previous large losses and therefore EWMA volatilities are relatively low.

Based on the low EWMA volatilities, the conditional Student t-distribution model only passes the tests in the recovery period, but appears to underestimate the risk measures in all periods. Furthermore, during the good economy period both *VaR* and *ES* are significantly underestimated when determined by conditional EVT models, hence this conditional EVT methods are rejected. Lastly, given the fact that the unconditional EVT models, which are

used to estimate both *VaR* and *ES*, pass all the tests during every period, it seems unnecessary to use the dynamic EVT, which will take current market condition into account.

Aside from similarities, there are some discrepancies between the best forecasting *VaR* models and the best estimating *ES* models. Even though the unconditional Normal distribution cannot be rejected during good economy period, it appears to overestimate *ES* but underestimate *VaR*. Similarly, all the EVT models in the recovery period and the unconditional EVT model with $\xi \neq 0$ in the good economy period show overestimated *ES*, but underestimated *VaR*. On the other hand, the unconditional Student t-distribution underestimates *ES* but overestimates *VaR* during the recovery period, despite the fact that this method is accepted for both *VaR* and *ES*, and the same goes to the unconditional EVT models and the unconditional EVT model with $\xi \neq 0$ during the bad economy period.

Furthermore, while the EVT model with $\xi = 0$ generates almost the same result as the corresponding EVT model with $\xi \neq 0$ in the same situation, when using as a *VaR* estimating tool, the *ES* is overestimated by the conditional EVT with $\xi = 0$, but not by the corresponding model with $\xi \neq 0$ during the financial crisis period. The same situation applies to the unconditional EVT models for the good economy period: the unconditional EVT model with $\xi \neq 0$ overestimates *ES*, while the unconditional normal EVT model underestimates it.

5 Summary and Conclusion

This paper is composed with an aim to compare the best models to calculate *VaR* and *ES*. Nine models are developed from four estimation approaches. Under the parametric approach, *VaR* and *ES* were estimated with the unconditional HS model. Under the Gaussian distribution approach, two models, namely the conditional and unconditional N-distribution were employed. Similar to the Gaussian distribution approach, two models, conditional and unconditional t-distribution, were used under the Student t-distribution approach. Lastly, under the Generalized Pareto distribution approach, four EVT models were applied, which were conditional EVT with $\xi = 0$ and $\xi \neq 0$, unconditional EVT with $\xi = 0$ and $\xi \neq 0$.

On account of the volatility clustering, EWMA model was applied to the conditional models. To implement the calculation of *VaR* and *ES* under EVT, the 95th percentile point of the empirical distribution was used as the threshold value and the parameters were updated for each evaluation window. To find the best models, the nine models were utilized to estimate both *VaR* and *ES* and then backtested with Kupiec test and the second method proposed by [Acerbi and Szekely \(2015\)](#) respectively.

To obtain a more accurate comparison of the models between the former and the latter risk measure, the analyses of the best models for each risk measure were first conducted separately.

The first part of Chapter 4 highlighted the performance analysis of the estimation models for *ES*. It was found that different estimation models and various testing periods led to varying the degree of estimation accuracy. Moreover, the HS model, the unconditional and conditional Normal distribution models along with the unconditional and conditional Student t-distribution models did not seem to be good choices for the overall period. On the other hand, the unconditional normal EVT models were accepted as good models for all time. Unexpectedly, the result also indicated that conditional models did not improve the accuracy of *ES* estimates compared to corresponding unconditional models, because the EWMA gave lesser weight to previous losses. EWMA lowered the volatilities, which led to

underestimation of risk measures under the conditional Normal distribution and Student t-distribution models.

The second part of Chapter 4 displayed the investigation of *VaR* accuracy under each models. The record showed that the HS model, the unconditional and conditional Normal distribution models along with the unconditional and conditional Student t-distribution models were rejected in at least one evaluation window. On the contrary, the unconditional EVT models produced reasonable estimates of *VaR* in any period. Additionally, conditional models worked even worse than corresponding unconditional models in the family of Normal distribution and Student t-distribution models.

After separate analysis for *VaR* and *ES* estimation models, these models were, the best models for both risk measure were finally compared. It was observed that the best performing *VaR* models were almost the same to the best performing *ES* models, which is similar to the findings of [Žiković and Dizdarević \(2011\)](#). The unconditional EVT models were the only models that passed the backtesting for every period. In other words, the unconditional EVT models seemed to be the only suitable models for all three evaluation periods. Furthermore, changing from *VaR* to *ES* reduce the probability of risk underestimation problem, which has been the main concern of the financial regulator.

Due to the several limitations of this research, there are several areas of improvements for the future research. Considering the failure of the current conditional models that capture the volatility clustering with EWMA approach, the future study could develop other conditional models with different approach to capture volatility clustering, such as GARCH or LM-GARCH. Moreover, future research could also be conducted with more non-parametric approaches and on different market aside from the S&P 500 index.

References

- Acerbi, C. & Szekely, B. (2014). Backtesting Expected Shortfall: Introducing three model-independent, non-parametric back-test methodologies for Expected Shortfall [pdf] Available at: <https://www.msci.com/documents/10199/faf1d99b-6c52-42cd-9b74-c2c95720b8b5>
- Acerbi, C. & Szekely, B. (2015). Backtesting Expected Shortfall [pdf] Available at: https://workspace.imperial.ac.uk/mathfin/Public/Seminars%202014-2015/Acerbi_January2015_Slides.pdf [Accessed 30 March 2015]
- Acerbi, C. & Tasche, D. (2001). Expected Shortfall: a natural coherent alternative to Value at Risk [pdf] Available at: <http://arxiv.org/pdf/cond-mat/0105191.pdf>
- Acerbi, C. & Tasche, D. (2002), Expected Shortfall, A Natural Coherent Alternative to Value at Risk, *Economics Notes*, 31(2): 1–10 [pdf] Available at: <http://onlinelibrary.wiley.com/doi/10.1111/1468-0300.00091/epdf>
- Al-Maghaireh, A. & Awartani, B. (2012): Modelling and Forecasting Value-At-Risk in the UAE Equity Markets: the role of long memory, fat tails and asymmetries in return innovations. *Review of Middle East Economics and Finance*, 8, pp. 1–22.
- Aloui, C. & Hamida, H., b. (2014). Modelling and forecasting value at risk and expected shortfall for GCC stock markets: Do long memory, structural breaks, asymmetry, and fat-tails matter?. *North American Journal of Economics and Finance*, 29 (2014) 349–380, Available at ScienceDirect
- Association for Financial Markets in Europe (n.d.). Definition of Capital. Available at: <http://www.afme.eu>
- Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1997). Thinking coherently, *Risk*, 10, vol. 11, pp. 68 – 71.

- Artzner, P., Delbaen, F., Eber, J.M., & Heath, D. (1999). Coherent Measures of Risk, *Mathematical Finance*, 9, (3), pp. 203–228.
- Baran, J. & Witzany, J. (2011). A Comparison of EVT and Standard *VaR* Estimations, SSRN, Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1768011
- Basel Committee on Banking Supervision. (1988). International Convergence of Capital Measurement and Capital Standards, Basel:BIS [pdf] Available at: <http://www.bis.org/>
- Basel Committee on Banking Supervision. (2015). An Explanatory Note on the Basel II IRB Risk Weight Functions, Basel:BIS [pdf] Available at: <http://www.bis.org/>
- Basel Committee on Banking Supervision. (2014). A Brief History of The Basel Committee, Basel:BIS [pdf] Available at: <http://www.bis.org/>
- Basel Committee on Banking Supervision. (2014). Fundamental review of the trading book: A revised market risk framework, Basel: BIS [pdf] Available at: <http://www.bis.org/>
- Bellini, F. & Bignozzi, V. (2014). Elicitable Risk Measures, *Quant. Finance*, Vol 15(5), pp. 725-733. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2334746
- Blattberg, R.C. & Gonedes, N.J. (1974). A Comparison Of The Stable And Student Distribution As Statistical Models For Stock Prices, *The Journal of Business*, Vol. 47, No. 2, pp. 244–280.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 31, pp. 307–327.
- Boudoukh, J., Richardson, M. & Whitelaw, R. (1998). The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk, *Risk*, 11, pp. 64–67.
- Danielsson, J. & de Vries C.G. (2000). Value-at-Risk and Extreme Returns, *Annales d'Economie et de Statistique*, ENSAE, 60, pp. 239–270.
- Degiannakis, S., Floros, C., & Dent, P. (2013). Forecasting value-at-risk and expected shortfall using fractionally integrated modelsof conditional volatility: International evidence. *International Review of Financial Analysis*, 27, 21–33. Available at: Lund University Libraries.

- Dowd, K. (2005). *Measuring Market Risk*, West Sussex: John Wiley & Sons Ltd.
- Echaust, K. & Just, M. (2013). Conditional Versus Unconditional Models For *VaR* Measurement. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2365588
- Fama, E. (1965). The Behavior Of Stock Market Prices, *The Journal of Business*, 38, pp. 34–105.
- Gençay, R. & Selçuk, F. (2004) Extreme Value Theory and Value-at-Risk: Relative Performance in Emerging Markets, *International Journal of Forecasting*, 20, pp. 287-303
- Gneiting, T. & Ranjan, R. (2011). Comparing Density Forecasts Using Threshold- And Quantile-Weighted Scoring Rules, *Journal of Economics and Business*, Stat. 29, pp. 411–422. Available at: Lund University Libraries.
- Gilli, M. & Këllezzi, E. (2006). An Application of Extreme Value Theory for Measuring Financial Risk, *Computational Economics*, 27(1), pp. 1–23.
- Harmantzis F., Miao, L. & Y. Chien. (2006). Empirical Study of Value-at-Risk and Expected Shortfall Models with Heavy Tails, *The Journal of Risk Finance*, Vol. 7(2), pp. 117–35.
- Hull, J. C. (2012). *Risk Management and Financial Institution*, 3rd edn, New Jersey: John Wiley & Sons, Inc.
- Hull, J. & White, A. (1998). Incorporating Volatility Updating into the Historical Simulation Method for Value at Risk, *Journal of Risk*, pp. 5–19.
- Hürlimann, W. (2004). Distortion Risk Measures And Economic Capital, *North America Actuarial J.*, 8, pp. 86–95.
- Hürlimann, W. (2008). Solvency II Reinsurance Counterparty Default Risk, *Life and Pensions Magazine*, December Issue, pp. 39–44.
- Hürlimann, W. (2009). Risk Management For Solvency II And Beyond: A survey of some recent methods. *White Paper FRSGlobal/Wolters Kluwer Financial Services*, Available at: http://www.solvency2explained.com/WP_RiskmanagementforSolvencyII070711.pdf .

- Jadhav, D., Ramanathan, T. V. & Naik-Nimbalkar, U.V. (2009) Modified Estimators of Expected Shortfall, *Journal of Emerging Market Finance*, 8:2, pp.87 – 107. Available at: Lund University Libraries
- Jalal, A. & Rockinger, M. (2008). Predicting Tail-Related Risk Measures: The consequences of using GARCH filters for non-GARCH data, *Journal of Empirical Finance*, 15, pp. 868–877.
- JP Morgan. (1996). Risk Metrics Monitor, Fourth Quarter, New York: Morgan Guaranty Trust Company.
- Kerkhof, J. & Melenberg, B. (2004). Backtesting for Risk-Based Regulatory Capital, *Journal of Banking & Finance*, vol. 28, no. 8, pp.1845-1865, Available Online: <http://www.sciencedirect.com/science/article/pii/S0378426603002255> [Accessed 13 April 2015]
- Koji, I. & Kijima, M. (2003). On the Significance of Expected Shortfall as a Coherent Risk Measure, *Journal of Banking & Finance*, Vol. 24(9), pp. 853–64.
- Kupiec, P. (1995). Techniques for Verifying the Accuracy of Risk Management Models, *Journal of Derivatives*, 3, pp. 73–84.
- Liang, B. & Park, H. (2007). Risk Measures For Hedge Funds: A cross-sectional approach, *Europe Financial Management*, 13, pp. 333–370.
- Mandelbrot, B. (1963). The Variation Of Certain Speculative Prices, *Journal of Business*, 36, pp. 394 - 419.
- Marinelli, C., D'Addona, S., & Rachev, S. (2007). A Comparison Of Some Univariate Models For Value-At-Risk And Expected Shortfall, *International Journal of Theoretical and Applied Finance*, 10, pp. 1043–1075.
- McNeil, A.J. & Frey, R. (2000). Estimation Of Tail-Related Risk Measures For Heteroscedastic Financial Time Series: An Extreme Value Approach, *Journal of Empirical Finance*, no. 7 (3-4), pp. 271–300.
- MSCI Inc. (2014). MSCI Demonstrates That Backtesting *Expected Shortfall* is Possible and Could Potentially Replace *Value at Risk (VaR)* in the Calculation of Regulatory Capital

Requirements: New methodology ends debate as to whether Expected Shortfall can be backtested [pdf] Available at: <https://www.msci.com/documents/10199/92848255-89e6-46e7-9522-2af00df50566>

Odening, M. & Hinrichs, J. (2003). Using Extreme Value Theory to Estimate Value-at-Risk [pdf] Available at: <http://www.agrar.huberlin.de/de/institut/departments/daoe/abl/dokumente/literatur/VAREVT/view>

Oh, S. & Moon, S., J. (2006). Comparative Analysis Of Portfolio Risk Measures Based On EVT-Copula Approach During Financial Crises, *Asia-Pacific Journal Financial Study*, 35, pp. 175–205.

Pattarathammas, S., Mokkhaveva, S. & Nilla – Or, P. (2008). Value-at-Risk and Expected Shortfall under Extreme Value Theory Framework: An Empirical Study on Asian Markets [pdf] Available at: <http://www.bus.tu.ac.th/>

Perignon, C. & Smith, D.R. (2010). The Level And Quality Of Value-At-Risk Disclosure By Commercial Banks, *Journal of Banking and Finance*, 34, (2), pp. 362 – 377.

Pritsker, M. (2006). The Hidden Dangers Of Historical Simulation, *Journal of Banking and Finance*, 30, pp. 561–582.

Rachev, S. & Mittnik, S. (2000). *Stable Paretian Models in Finance*, New York: John Wiley & Sons

Reserve Bank of New Zealand (2007). Capital Adequacy Ratios For Banks - Simplified Explanation And Example Of Calculation [pdf] Available at: http://people.stern.nyu.edu/igiddy/articles/capital_adequacy_calculation.pdf

Righi, M., B. & Ceretta, P. S. (2015). A Comparison Of Expected Shortfall Estimation Models, *Journal of Economics and Business*, 78, pp. 14–47.

Saigol, L. & Fleming S. (2013). What the Basel Proposals on Bank Trading-Book Rules Mean, *Financial Times*, 31 October, Available Online: <http://www.ft.com/intl/cms/s/0/e4d4a130-421d-11e3-9d3c-00144feabdc0.html#axzz3WvIMKkBX/>[Accessed 30 March 2015]

- Singh, A., K., Allen, D., E. & Robert, P., J. (2013). Extreme Market Risk And Extreme Value Theory, *Mathematics and Computers in Simulation*, 94, pp. 310–328
- Song X. C. (2008), Nonparametric Estimation of Expected Shortfall, *Journal of Financial Economics*, no. 6, pp. 87–107 [pdf] Available at: <http://www.public.iastate.edu/~songchen/shortfall.pdf>
- Tolikas, K. (2014). Unexpected Tails In Risk Measurement: Some international evidence, *Journal of Banking and Finance*, 40, pp. 476–493.
- Wong, W., K., Fan, G., & Zeng, Y. (2012). Capturing Tail Risks Beyond *VaR* [pdf] Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1263949
- Yamai, Y. & Yoshiba, T. (2002). Comparative Analyses Of Expected Shortfall And Value-At-Risk: their estimation error, decomposition, and optimization, *Monetary Economy Study*, pp. 87–122.
- Yamai, Y. & Yoshiba, T. (2005). Value-at-Risk Versus Expected Shortfall: A Practical Perspective, *Journal of Banking & Finance*, 29, pp. 997–1015, Available at: www.sciencedirect.com
- Ziegel, J., F. (2014). Coherence and Elicibility [pdf] Available at: <http://arxiv.org/abs/1303.1690v3>
- Žiković, S. & Vlahinić –Dizdarević, N. (2011). Similarities between Expected Shortfall and Value at Risk: Application to Energy Markets, *International Journal of Management Cases*, Vol. 13 Issue 3, pp. 386 Available at: Lund University Libraries.

Appendix A

Generalized Pareto Distribution (GPD)

When $\xi \neq 0$:

$$\begin{aligned}
 ES_\alpha &= \frac{1}{1-\alpha} \int_\alpha^1 \left\{ u + \frac{\beta}{\xi} \left[\left(\frac{N}{N_u} (1-x) \right)^{-\xi} - 1 \right] \right\} dx \\
 &= \frac{1}{1-\alpha} \int_\alpha^1 \left(u - \frac{\beta}{\xi} \right) dx + \frac{1}{1-\alpha} \int_\alpha^1 \frac{\beta}{\xi} \left(\frac{N}{N_u} \right)^{-\xi} (1-x)^{-\xi} dx \\
 &= \frac{1}{1-\alpha} \left(u - \frac{\beta}{\xi} \right) x \Big|_\alpha^1 + \frac{1}{1-\alpha} \frac{\beta}{\xi} \left(\frac{N}{N_u} \right)^{-\xi} \frac{1}{\xi-1} (1-x)^{-\xi+1} \Big|_\alpha^1 \\
 &= u - \frac{\beta}{\xi} - \frac{1}{1-\alpha} \frac{\beta}{\xi} \left(\frac{N}{N_u} \right)^{-\xi} \frac{1}{\xi-1} (1-\alpha)^{-\xi+1} \\
 &= u - \frac{\beta}{\xi} + \frac{1}{1-\xi} \frac{1-\alpha}{1-\alpha} \frac{\beta}{\xi} \left(\frac{N}{N_u} \right)^{-\xi} (1-\alpha)^{-\xi} \\
 &= u - \frac{u}{1-\xi} - \frac{\beta}{\xi} + \frac{1}{1-\xi} \frac{\beta}{\xi} + \frac{1}{1-\xi} \left[u + \frac{\beta}{\xi} \left(\frac{N}{N_u} \right)^{-\xi} (1-\alpha)^{-\xi} - \frac{\beta}{\xi} \right] \\
 &= \frac{\beta - u\xi}{1-\xi} + \frac{VaR_\alpha}{1-\xi}
 \end{aligned}$$

When $\xi = 0$:

$$\begin{aligned}
 ES_\alpha &= \frac{1}{1-\alpha} \int_\alpha^1 \left[u - \beta \ln \left(\frac{N}{N_u} (1-x) \right) \right] dx \\
 &= \frac{1}{1-\alpha} \left[u - \beta \ln \left(\frac{N}{N_u} \right) \right] x \Big|_\alpha^1 + \frac{1}{1-\alpha} \int_\alpha^1 [-\beta \ln(1-x)] dx \\
 &= u - \beta \ln \left(\frac{N}{N_u} \right) - \beta [\ln(1-\alpha) - 1] \\
 &= u - \beta \ln \left(\frac{N}{N_u} (1-\alpha) \right) + \beta \\
 &= VaR_\alpha + \beta
 \end{aligned}$$

Appendix B

Table 5: Parameters estimates for Extreme Value Theory

		EVT		EVT-cond	
		$\xi = 0$	$\xi \neq 0$	$\xi = 0$	$\xi \neq 0$
Bad Economy Period	ξ	0	-0,1115	0	0,2699
	β	2,0905	2,3318	0,7320	0,5188
	$q_{0,99}$	-	-	3,2372	3,1071
	$q_{0,975}$	-	-	2,5665	2,4513
Recovery Period	ξ	0	-0,1577	0,0000	0,0654
	β	1,1154	1,2924	0,6406	0,5993
	$q_{0,99}$	-	-	2,9603	2,9464
	$q_{0,975}$	-	-	2,3733	2,3541
Good Economy Period	ξ	0	-0,4026	0	-0,4387
	β	0,4471	0,6476	0,4490	0,6706
	$q_{0,99}$	-	-	2,0626	2,1132
	$q_{0,975}$	-	-	1,6512	1,7411

Note: The threshold value u is equal to the 95th percentile point of the empirical distribution. The parameters are updated for each period.

Appendix C: Part A

Expected Shortfall at 97.5% Confidence Level:



Figure 2: Expected Shortfall - HS

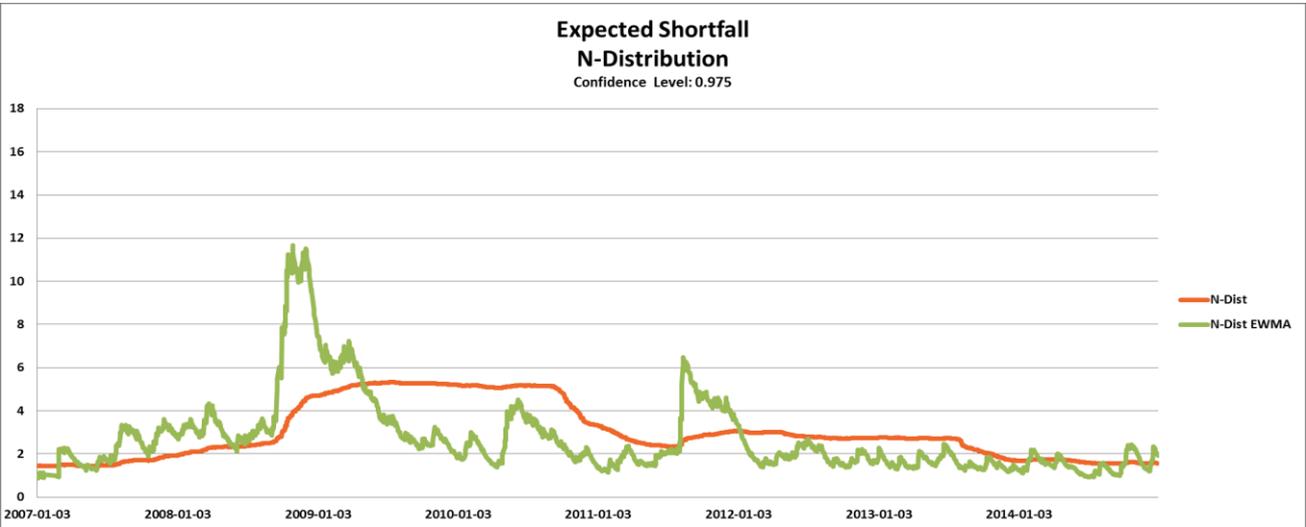


Figure 3: Expected Shortfall - N-dist

Appendix C: Part B

Expected Shortfall at 97.5% Confidence Level:

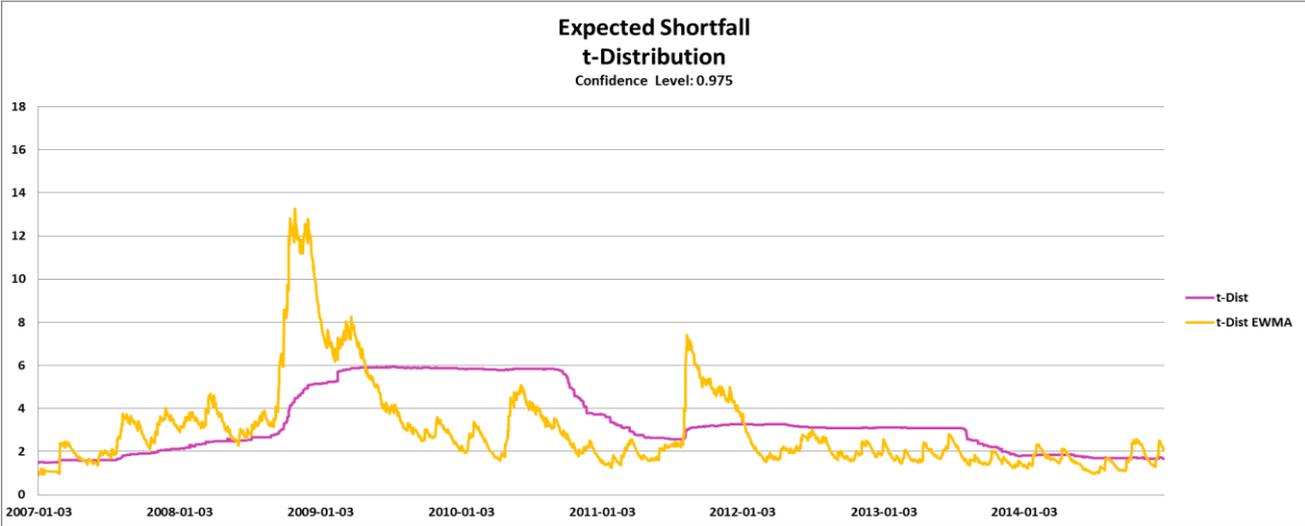


Figure 4: Expected Shortfall – t-dist



Figure 5: Expected Shortfall - Unconditional EVT

Appendix C: Part C

Expected Shortfall at 97.5% Confidence Level:

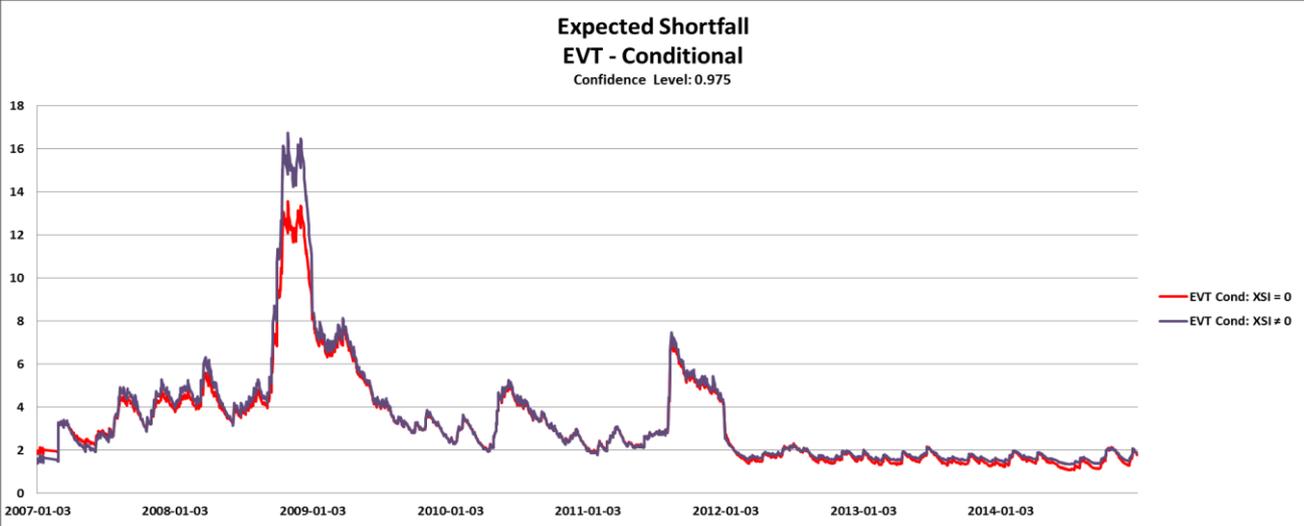


Figure 6: Expected Shortfall - Conditional EVT

Appendix D: Part A

Value – at – Risk at 99% Confidence Level:

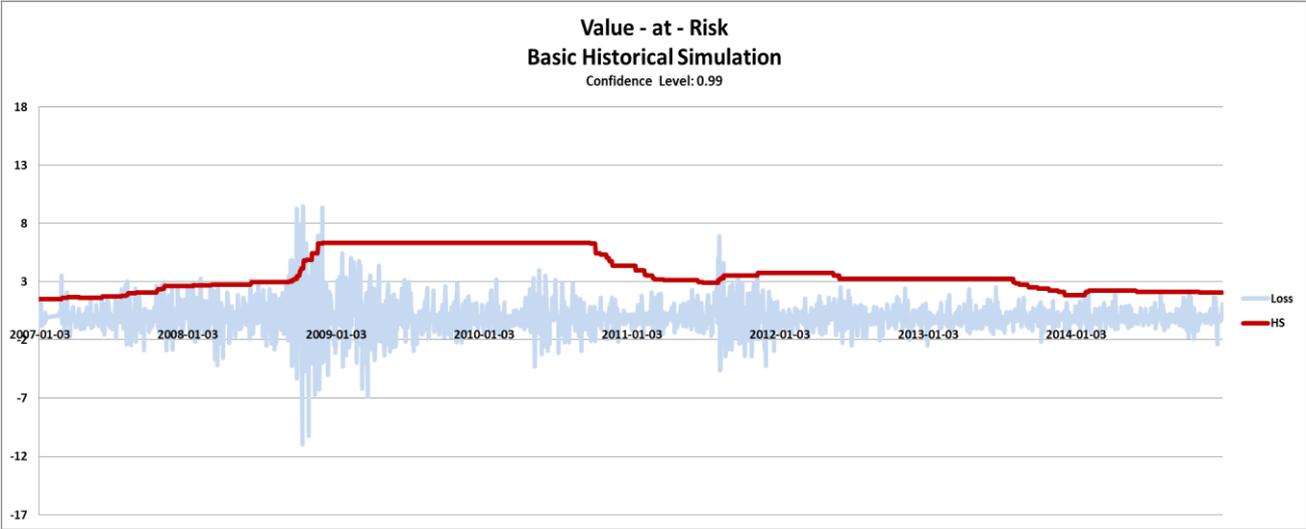


Figure 7: Value at Risk - HS



Figure 8: Value at Risk - N-dist

Appendix D: Part B

Value – at – Risk at 99% Confidence Level:



Figure 9: Value at Risk - t-dist

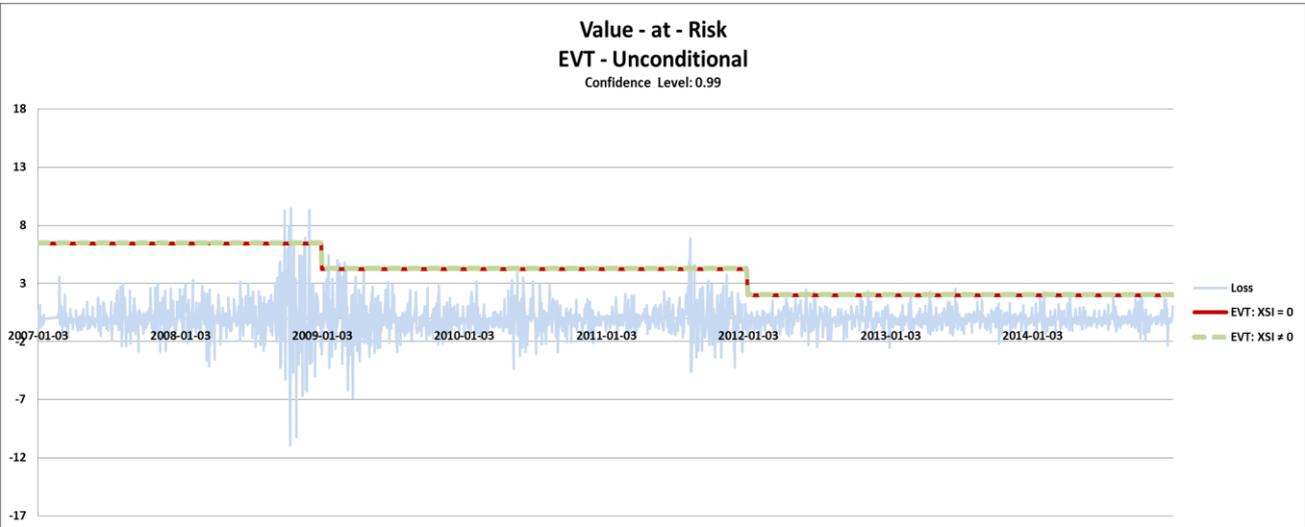


Figure 10: Value at Risk - Unconditional EVT

Appendix D: Part C

Value – at – Risk at 99% Confidence Level:



Figure 11: Value at Risk - Conditional EVT